ENGINEERING AFFIBODIES TO INHIBIT GLUCOSYLTRANSFERASES USING YEAST DISPLAY

by

Angel D'Oliviera

A thesis written in the partial satisfaction of the requirements for the degree of Master of Science in Biotechnology

Baltimore, MD

August 2017

I. Abstract

Dextransucrases are a type of Glucosyltransferase (Gtf) which have many industrial applications, including oral care and baking, and have been identified as a biologically relevant target to inhibit. There are many advantages to designing a small protein affibody to accomplish this goal compared to using antibodies and small molecule inhibitors, including their inherent stability, expressibility, and engineerability. Affibodies are small protein domains, the original IgG binding surface of which can be varied combinatorially to produce tight binders to diverse protein targets. Yeast display provides a powerful tool to assemble and rapidly screen large libraries for binders against the target of interest, Gtfs. Initial studies showed the wild-type affibody could be successfully expressed and exhibit binding to an IgG antibody in the yeast display system. Following this, a valid approach to designing, assembling, and expressing an affibody variant library in this yeast display system was systematically employed. The assembled library contained an estimated 2.5 x 10^8 members, and a sampling of this library illustrated the high quality and diversity of these variants. Initial surface display analysis indicated sufficient library expression level, as well as an anticipated disruption to IgG binding.

The affibody variant library was enriched and sorted for Gtf binders using fluorescent activated cell sorting under several different conditions. Upon analysis of the binding populations, a potential Gtf binding motif was uncovered. Moreover, a dominant variant identified was found to be present in the majority of the libraries analyzed, regardless of sorting conditions. This highly represented variant was confirmed as a specific Gtf binder, with a dissociation constant in the range of tens to hundreds nanomolar. In addition to this validated binder, fifteen variants of interest were expressed and analyzed

for Gtf inhibition. Though no Gtf inhibitors were found in these select few variants analyzed, the results from this screening demonstrate great promise to obtain such a molecule. Furthermore, the findings from this library screening could be strategically leveraged in the future using an optimized approach to maximize chances of identifying a Gtf inhibitor.

Thesis Readers: Dr. Josh Sims, Dr. Kristina Obom, Dr. Robert Lessick
Acknowlegements

I would like to thank my incredible mentor, Dr. Josh Sims, for the tremendous advice, guidence, and motivation provided during this project. Every last-minute meeting, critical discussion, and rapidly answered weekend email was invaluable in driving this research. I could not have asked for a more enthusiatic and contemplative mentor (plus no one else appreciated my terrible science puns), and I have become a better scientist because of this experience. Special thanks to Jon Peterson helping me set up and run the FACS, and aiding during complex brainstorming sessions.

I have much gratitude for my manager, Tanja Gruber for fully encouraging this academic endeveaor, and to DuPont Industrial Biosciences for funding this effort. Dr. Kristina Obom and Dr. Thomas Koval also provided helpful direction and encouragement throughout the duration of this project.

I could not have gotten through my program and this research without the support of my fantastic friends and family, in addition to the unwavering patience and emotional backing of my roomate. Finally, a big thanks to Hartsy, who was an unconditionally loving companion during my literature searches and paper writing.

Table of Contents

List of Tables

List of Figures

II. Introduction

Glucosyltransferases

Glucosyltransferases (Gtfs) are a well-known class of enzymes which have the ability to transfer a sugar residue from a donor to an acceptor molecule, forming glycosidic linkages. Bacterial Gtfs are vital to the biosynthesis of peptidoglycan, which provides cell wall rigidity (1, 2). Gtf members of the glycoside hydrolase family 70, known as dextransucrases, primarily create soluble alpha 1,6-linked polymers from sucrose (**Figure 1**). Dextransucrases are specifically found in seven species of the lactic acid bacteria genera: *Leuconostoc, Lactobacillus, Streptococcus, and Weissella* (3). *Weissella confusa* and *Weissella cibaria* dextransucrases are distinct enzymes with application in sourdough fermentation. The addition of these dextran-producing strains can improve the texture and quality of bread (3, 4).

Gtfs, including dextransucrases, from *Leuconostoc* and *Streptococcus* have been implicated as a driving force behind the formation of plaque (biofilms) (5); through the secretion of these enzymes, sucrose is polymerized into an extracellular structural matrix to allow the formation of a biofilm of entire communities of oral bacteria on enamel (1, 6). Secreted dextransucrases can adsorb to the tooth surface and produce soluble material, which can be used as a primer for primer-dependent Gtfs to generate insoluble polysaccharides (primarily alpha 1,3-glucosyl linkages), which is necessary to provide sites for matrix formation (1, 2). Approximately a fifth of the dry weight of plaque is composed of polysaccharides (7).

Due to their diverse functionality and uses, Gtfs have been recognized as an industrial relevant target in a variety of applications. Identification of inhibitors to these enzymes could be beneficial in novel ways. For instance, a Gtf inhibitor could serve as an antimicrobial molecule with potential to be used in the oral hygiene market (8). Additionally, an inhibitor could provide a way to adjust the structure of dextran being generated (7) which could provide advantageous new qualities in the food industry.

Glucosyltransferase Inhibitors

Targeted inhibition of Streptococcal Gtf enzymes may provide an effective antimicrobial approach; however, this is a challenging endeavor because the donor and acceptor molecule binding affinities are unusually low (9). Additionally, small inhibitor molecules may result in acquired bacterial resistance, which is difficult to combat with small molecules (10). There are several Gtf-specific antibodies available today. However, the binding of a Gtf antibody does not necessarily inhibit activity, and of those which truly inhibit activity, there are very few which show efficacy *in vivo* (11). It is also important to note that antibody discovery and production are typically expensive ventures. Plaque presentation requires continuous dosing, making low production costs of an inhibitor a desirable objective. Additionally, the large molecular size of antibody inhibitors may limit diffusive penetration to certain niches.

Numerous Gtf inhibitors to both crude Gtf extracts and specific Streptococcal enzymes have been previously found, including over thirty inhibitors which were derived from natural products, synthesized inhibitors, antibodies, metal ions and oxidizing agents. Of these inhibitors, only four natural products, one synthesized inhibitor, and a single

antibody have had evidence of cariostatic efficacy *in vivo*. The major challenge acknowledged in this pursuit is the ability to inhibit the Gtf enzymes which have adhered to the surface of the enamel or are attached to the surface of the bacterial cell. Furthermore, none of these molecules have yet been cleared for therapeutic use. (11-13) The difficulty of identifying Gtf inhibitors is attributed to their complicated reaction mechanisms, as well as their unusual conformational plasticity (14). Inhibitors do not necessarily need to sterically block the active site of the enzyme, but can also interfere with the catalytic mechanism of the enzymes (14, 15). Presumably, disruption of the folding or supra-molecular structure of Gtfs may also result in effective inhibition.

Interestingly, it has been found that some Gtf antibodies can influence not only the amount of Gtf polysaccharide production, but also the shape of the material. This concept was found to be particularly relevant for Streptococcal Gtfs, as primer-dependent Gtfs could have a decreased ability to bind to Gtf-priming material generated in the presence of the Gtf antibody. The shape of the enamel matrix formation can be affected, ultimately altering the formation of biofilms. Additionally, the presence of Gtf antibodies was found to modify the actual sugar linkages themselves; Gtfs which formed primarily insoluble alpha 1,3-linkages produced an increased amount of soluble alpha 1,6-linkages in the presence of a Gtf antibody. The ability to change the quantity, linkage, and solubility of the material could be beneficial in industrial applications while also providing new functional understanding about these complex enzymes. (7)

Affibodies

A small, economically produced protein scaffold could not only provide a way to inhibit Gtf enzymes, but also offer a plasticity to predict and address future bacterial resistance. Affibodies are significantly smaller than antibodies, and will also have the advantage of stability and ease of production. This 58-amino acid protein was derived from a staphylococcal cell wall protein. There are no disulfide bonds in affibodies, which contain three alpha helices (**Figure 2**) that can be reversibly folded. Affibodies have been used as engineerable scaffolds for many different biotechnological, diagnostic, and therapeutic applications (16). Thirteen solvent-exposed residues on helix one and two, originally an IgG-binding domain, can be randomized to acquire new binding properties (16, 17). The small size, high binding specificity, lack of cross-reactivity, high binding affinity, and overall stability properties make affibodies optimal proteins to engineer for new binding activities (18).

Affibodies have been successfully targeted to a wide variety of proteins. For instance, affibodies have been designed as enhanced picomolar affinity binders to epidermal growth factor receptor 2, a cancer specific cell surface receptor (16). Many other tumor-targeting antigens have been targeted successfully by affibody libraries, including EGFR, IGF1R, and HER3 in the low nanomolar affinity range. Affibodies have also been designed as biosensors, imaging molecules, and capture agents. (17-19)

Yeast Display

Protein display systems are powerful protein engineering tools used for rapid *in vitro* protein evolution. In these systems, bacterial, fungal, and viral hosts can be used to

express library variants on the surface of the cells. Fluorescence activated flow cytometry (FACS) is used in conjunction with display systems to successfully isolate binders with high affinity to new targets. *Saccharomyces cerevisiae* has been identified as a particularly advantageous surface display host in which to assemble large numbers of well-folded and highly-expressed variants (20).

Using an inherent ability called gap repair, recombinant DNA can be rapidly assembled within yeast cells. This process allows linearized DNA fragments, typically containing a digested replication vector and the systematically randomized sequence of the variant, to be transformed into the yeast where they are assembled into circular DNA by the cell DNA repair machinery. In addition to this effective transformation method, the yeast cells also contain a strict protein quality control system, minimizing the generation of misfolded or otherwise incorrect protein variants. (20)

To achieve yeast surface expression of the protein scaffold of interest, the affibody variant is fused to the C-terminus of the yeast protein Aga2 (**Figure 3A**). The addition of a C-terminal epitope tag allows expression and binding to be monitored independently. Aga2 is a mating protein which covalently binds to Aga1, a GPI-anchored membrane-associated protein on the yeast surface. In this yeast display system, these mating proteins are regulated by a galactose inducible promoter. When using this system to express variants, this inducibility allows for the decoupling of the expression of the engineerable protein and the cellular division. This is desirable for variants which may be toxic to the cell, and is a property which provides a benefit over other display systems (i.e. phage display). The final induced yeast cell displays a variant anchored on the cell surface and capable of interacting with extracellular molecules in the cell suspension (**Figure 3B**).

Since the plasmid system used to harbor the Aga2-fused variant has a centromeric origin of replication, the copy number maintained in the cell is essentially one, allowing for efficient, 1:1 coupling of the genotype and phenotype. (20, 21)

Using an affibody protein as a scaffold backbone, large libraries of variants can be displayed on the surface of yeast and screened for target binding using fluorescence activated cell sorting (FACS). Flow cytometric sorting of *S. cerevisiae* surface display libraries is a well-established technique to separate desired cells from a complex mixture (**Figure 4**). Yeast cells expressing library variants on their surface can be incubated with a fluorescently-labeled target molecule which can bind to specific variants. This renders only the cells containing variants with target binding properties fluorescent. An additional antibody with affinity for the C-terminal epitope tag and a different fluorescent emission wavelength, can be used to detect the overall expression of the affibody. A flow cytometer can detect and sort the cells which emit a signal from those without fluorescence.

With this method, yeast cells with highly expressed variants, as well as target binding activity, can be selected. The separated cells can be recovered and sequenced to identify the variant with binding activity. Millions of individual cells can be analyzed and sorted in a relatively short amount of time, providing an exceptional platform to screen large libraries of variants within a single mixture (20). Often, a weak binder to a target is obtained from a large and diverse library after only a few rounds of sorting. Further mutagenesis can result in designing a tight binder to the target, which can then be tested for biological inhibition.

III. Materials and Methods

Gtf Purification

Previously generated *E. coli* ultrafiltrate of recombinant Dextransucrase from *W. confusa* (AHU88292) was obtained from DuPont Industrial Biosciences (Dr. Rong Guan). Sample was diluted in half with 2x binding buffer (40mM Sodium Acetate, pH 5.5). 2% bed volume (v/v) of a 50% slurry of preequilibrated Sephadex75 resin was added to sample and rocked for 30 mins at room temperature. The resin was applied to a 10ml benchtop column and washed with 10x column volumes (CV) of binding buffer. Protein was eluted with 1% dextran (MW 10000). Sample was concentrated using PES 3kD MWCO filtration and quantitated using A_{280}.

Gtf Labeling

Purified protein was buffer exchanged to PBS, pH 8.0 using a PES 3kD MWCO spin column. Biotinylation of purified Gtf was performed using the Solulink ChromaLink Biotin Protein Labeling Kit. Biotinylation was performed three separate times. Final molar ratio of Biotin:Gtf was 7.5, 15, 12 for runs 1, 2, and 3, respectively. FITC labeling of purified Gtf was performed using the Pierce™ FITC Antibody Labeling Kit. Final molar ratio of FITC:Gtf was 2.5.

Gtf Activity Assessment

Polymerization reactions were carried out in 20mM sodium acetate pH 5.5, 1% (v/v) of crude enzyme lysate, and 200g/L sucrose. Samples were incubated with agitation

at 40C for 18 hours. For quick analysis, sample was checked visually for high viscosity, manifested in a gel-like form. For more thorough analysis, activity was monitored using HPLC, measuring sucrose consumption over time.

Affibody Library Design

The Z-domain of staphylococcal protein A (pdb: 2SPZ) was used as a backbone for mutagenesis. Thirteen amino acids were targeted for an oligonucleotide-directed mutagenesis library (22), highlighted in cyan below:

1 10 20 30 40 50

VDNKFNKEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKLND

58

AQAPK

Codons were optimized for *Saccharomyces cerevisiae* using Integrated DNA Technologies Codon Optimization Tool. A Trimer Oligonucleotide Library was ordered from the Yale Keck Oligonucleotide Synthesis facility. These libraries use a mix of 19 codon trimer phosphoramidites to randomize selected amino acid positions, yielding a more uniform distribution of amino acids across the variable region than single base randomization schemes, and allowing for the exclusion of undesirable variation (here, cysteines and stop codons) that reduce library performance. The following synthesized single stranded oligonucleotides were ordered with the variable region represented by "xxx":

GGT GGA GTT GAT AAC AAG TTC AAC AAG GAA xxx xxx xxx GCG xxx xxx GAA ATT xxx xxx TTG CCC AAT TTG AAC xxx xxx CAG xxx xxx GCC TTC ATC xxx TCA CTG xxx GAC GAT CCA TCT CAA TCC GCG AAT

Oligonucleotide primers to the underlined region were designed as follows:

AfflibF':

GGATCCGGAGGTAGCGGATCAGGAGGTGGAGGCTCCGGAGGAGGTGGAGTT GATAACAAGTTCAACAAGGAA

AfflibR1':

AGAAGATCCCTTAGGAGCCTGTGCATCGTTTAATTTCTTGGCCTCGGCCAGTA AATTCGCGGATTGAGATGGATCGT

AfflibR2':

TCGTATGGGTATCCAGTAGACTTAGATTCAGATCCAGATCCAGAAGATCCCTT AGGAGCCTGT

Affibody Library Extension and Amplification

The single stranded oligonucleotides were first annealed by combining 10uM synthesized trimer library oligonucleotides with 10uM AfflibR' and 1X NEB buffer 2. Sample was incubated at 95C for 2 mins, then cooled to 25 C at 0.5 degree per minute. The

annealed oligonucleotides were then extended to double stranded DNA using the NEB DNA Polymerase I, Large (Klenow) Fragment in multiple 63ul reactions at 2.5U/reaction, 2.4uM annealed oligonucleotides, 1X NEB buffer 2, and 159uM dNTPs. Reaction was incubated at room temperature for 1 hour. Extension success and fragment concentration estimation were executed by ethidium agarose gel. The extended library was amplified using PCR and PFU ultra II polymerase.

Affibody Library Transfection into Saccharomyces cerevisiae

Large scale transfection was performed according to the Gietz High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method (23). Diversity was estimated by performing a series of dilutions of the transfected material in triplicate, plating on 100 x 15 mm petri dishes with SD-Trp Agar. The number of colonies on the 10/1,000,000 and 1/1,000,000 dilution plates were counted. After transfection, several passages of cells encompassing 10X library diversity was performed, assuming 1 OD_{600} = 18.5 million cells. A freezer stock encompassing 10X library members was made and stored at -80C.

Quality Control of Affibody Library

After three passages of the transfected culture, 6ml was prepped using a Zymoprep Yeast Miniprep Kit. The extracted DNA was transformed into NEB® 5-alpha Competent *E. coli* (High Efficiency) cells. 95 colonies were selected for ampicillin resistance and their DNA was prepared for Sanger sequencing using an Illustra TempliPhi DNA Amplification

Kit. Invotrogen's commercially available CYC1 transcription termination signal reverse primer was used to sequence affibody sequence.

Naïve Library Induction and Analysis

Yeast induction was performed using the Wittrup lab protocol (24). After induction, 1 million cell aliquots were collected and washed extensively with 1x PBS with 0.5% BSA. An aliquot was incubated with unlabeled NFKappaB p65 rabbit antibody at 1:50 dilution for 1 hour at room temperature. Cells were then collected and washed extensively, and then incubated with 1:50 dilution of anti-rabbit IgG Alexa 647 antibody for 15 minutes. Cells were collected and washed extensively. A 1:100 dilution of HA rabbit 488 antibody was added to cells. One aliquot was kept as an unstained control, while another was only stained with the HA antibody. Cell aliquots were analyzed with a NovoCyte Benchtop Flow Cytometer.

Bead Enrichment

Enrichment of the affibody library for Gtf binders was performed by overnight incubation of 10X library diversity with 100nM biotin-labeled Gtf at 4C on an inverter. Separation of Gtf bound cells was performed using Miltenyl Biotech MACS separation protocol. Briefly, incubated cells were washed extensively to remove free biotin-labeled Gtf. Streptavidin coated microbeads were added to the washed cells. Cell suspension was loaded onto an equilibrated MACS LS column placed in a MACS Separator magnetic field. The captured cells were washed extensively, then removed from the magnetic field to elute. Cells were regrown and induced for subsequent sorting on FACS.

FACS Sort1-Sort 3

The induced naïve affibody library and bead enriched library were incubated with 50nM FITC-labeled Gtf and Alexa 647 conjugated HA-Tag antibody for 2 hours on an inverter at RT. The resultant samples were sorted using FACS, selecting for high signal in both channels. Samples were regrown, re-induced, and analyzed for binding using a titration of Gtf concentrations. For this analysis, various titrations of Gtf were incubated with 20 million cell aliquots overnight at 4C on an inverter. An additional target, TNF-alpha was also included. Libraries were subsequently taken forward two additional FACS sorts using the described regrowth, induction, and overnight Gtf incubation method, with a target concentration of 6nM FITC-labeled Gtf.

Library DNA Extraction and Analysis

Library DNA was prepped using a Zymoprep Yeast Miniprep Kit. The extracted DNA was transformed into NEB® 5-alpha ElectroCompetent *E. coli* (High Efficiency) cells under ampicillin selection. DNA from 96 colonies was prepared for Sanger sequencing using an Illustra TempliPhi DNA Amplification Kit. Invotrogen's commercially available CYC1 transcription termination signal reverse primer was used to sequence affibody sequence. Sequence analysis was performed using Geneious 10.2.2.(25). Sequence logo was generated using WEBLOGO Version 2.8.2. (26, 27).

Affibody Variant Expression

Identified candidate sequences were codon optimized for *E. coli* using Integrated DNA Technologies Codon Optimization Tool and ordered as synthetic gene blocks. Each sequence was designed to have 25 base pairs of homology to the pET28a replicating plasmid including an N-terminal His-Tag. Each variant was assembled using NEBuilder® HiFi DNA Assembly Cloning Kit and transformed into NEB® 5-alpha Competent *E. coli* (High Efficiency) cells. DNA was prepared for Sanger sequencing using an Illustra TempliPhi DNA Amplification Kit. Sequence verified plasmids were transformed and expressed in One Shot™ BL21(DE3) Star Chemically Competent *E. coli*. Variants were purified using nickel affinity chromatography and analyzed using SDS-PAGE. Extinction coefficients were calculated using Expasy (28) and concentration was estimated using A_{280}.

Affibody Binding Validation

Purified Variant A was analyzed using an Octet HTX System from ForteBio. Using Anti-Penta-HIS Biosensors, Variant A was captured and immobilized. The binding curves were assessed using two-fold serial dilutions of purified Gtf. All analysis was performed in PBS with 0.5% BSA. The Octet software was used for binding analysis and dissociation estimation. Raw data was inverted for binding analysis fit curves.

Affibody Inhibition Assessment

Each purified affibody was pre-incubated with 100nM unlabeled Gtf for 1-2 hours at room temperature. Affibodies were dosed at approximately 20X the Gtf molar concentration. Pre-incubation was carried out in reaction buffer, 20mM sodium acetate, pH

5.5. Polymerization reactions were carried out in 20mM sodium acetate pH 5.5, 100g/L sucrose. Samples were incubated with agitation at 30C for 18 hours. For quick analysis, sample was checked visually for high viscosity, manifested in a gel-like form.

IV. Results and Discussion

Library Design, Assembly, and Fitness Assessment

To assess the feasibility of using the yeast display system for the expression of an affibody library, the Z-domain of staphylococcal protein A wild-type (Wt) affibody was first expressed using this technology and evaluated by flow cytometry (**Figure 5**). The Wt expression levels on the surface of the yeast after induction was quite high (approximately 75% of library), approaching the practical limits of this system (20). Furthermore, this protein appears to be properly folded in this system, as illustrated by the ability to detect binding to IgG antibodies.

After the successful expression of Wt affibody, an affibody variant yeast display library was constructed. Thirteen amino acids on two of the three alpha helices were selected for saturation mutagenesis (**Figure 6A**). The creation of this library was strategized to anneal, extend, and amplify synthetic single-stranded DNA containing randomized regions, and utilize inherent gap repair within the *S. cerevisiae* host to assemble the amplified region into a yeast display vector (**Figure 6B**). In each step, the annealed, Klenow, and PCR product were detected at the appropriate size by ethidium bromide stained agarose gel (**Figure 6C**). After large scale transfection (**Figure 7A**), the estimated diversity of the library members was 2.5×10^8.

Assessing the fitness of the library is critical to ensure proper diversity and confirm a lack of bias, maximizing the potential of acquiring variants with new binding affinities. For this analysis, DNA was extracted from a pool of the transfected material and transformed into *E. coli*. Ninety-five colonies were sequenced, and the residues at each of

the 13 varied positions were evaluated (**Figure 7b**). Nearly 80% of the library sampling contained proper in-frame sequences, with mutations located only at the intended positions. From these samples, the percent representation of each of the 20 naturally occurring amino acids inspected and compared to the expected representation. No obvious bias was observed for any amino acid residue. The lack of cysteine residues was to be expected, as the library mutation regions were synthesized without cysteine codons.

The initial surface expression of the assembled library was evaluated using flow cytometry (**Figure 8**). Approximately half of the library was properly expressed. As anticipated, a disruption to the natural binding to IgG was observed for the majority of the library population.

Identification and Preparation of Target Gtf

Originally, Streptococcal Gtf targets were selected for recombinant expression and purification. However, this endeavor required complex optimization and was abandoned due to time constraints. A Dextransucrase from *Weissella confusa* was previously expressed recombinantly in *E. coli*. This protein was purified **(Figure 9)** and labeled with biotin or Fluorescein isothiocyanate (FITC).

Library Enrichment and Fluorescence Activated Cell Sorting

Two approaches were employed to probe the affibody variant library for binders; The use of an "Enriched" library, which utilized a magnetic bead enrichment step prior to the first FACS sort, and a "Non-Enriched" library, which consisted of a small sampling of the naïve library void of any preselection steps before FACS sorting. To obtain the enriched

library, the biotinylated Gtf was incubated with the inducted naïve library at 10X the estimated diversity. The cells bound to the biotinylated target were separated with streptavidin coated magnetic beads and regrown and induced. Approximately 5-10 million cells were isolated with this method.

The Enriched and Non-Enriched library were incubated with 50nm of FITC labeled Gtf and HA antibody conjugated with Alexa 647 for 2 hours. The resulting samples were sorted using FACS, selecting for cells with both high Alexa 647 and FITC signal. Roughly 35 million events were sorted of the Enriched library, resulting in a selection of nearly 10,000 cells. For the Non-Enriched library, approximately 111 million events were sorting, selecting around 12,000 cells. An analysis of this first sort (**Figure 10**) illustrates a successful selection of binders which exhibit both high yeast surface expression and affibody binding signal, compared to the Naïve library.

To ensure the binders selected in this first step exhibited target specificity, these libraries were also tested for binding against an unrelated target, TNF-alpha conjugated with Alexa Fluor 488 (**Figure 11**). The emerging population of binders in quadrant 2 vanish against this target, fully illustrating specificity for the Gtf target.

Each library, Enriched and Non-enriched, went through a total of three sorts. Sort 2 and sort 3 utilized more stringent conditions, including a lower target concentration and a selection gate requiring higher binding signal. The purification of each library population can be seen after each sort with distinct populations emerging at low Gtf concentrations (**Figure 12**).

Identification of Variant Members within Sorted Libraries

Six libraries were selected for sequence analysis:

1. Non-Enriched Sort 2 (3nM Gtf)
2. Non-Enriched Sort 2 (6nM Gtf)
3. Enriched Sort 2 (6nM Gtf)
4. Non-Enriched Sort 3 Gate A (6nM Gtf)
5. Non-Enriched Sort 3 Gate B (6nM Gtf)
6. Enriched Sort 3 (6nM Gtf)

Sequences from 96 individual clones for each library were analyzed (**Table 1**). A total of 25 sequences were found in at least two libraries, which represents 21% of the total unique sequences (**Table 2**). It was of note to identify the increase of representation of several sequences between sort 2 and sort 3 of each library, illustrating the path of purification during each round.

For each of the libraries after sort 3, many unique sequences emerged (**Figure 13a**). The Non-Enriched Sort 3 Gate A had three unique sequences, the dominant making up 97% of the library, with a second sequence differing only by a single amino acid. This sequence was found in all of the libraries assessed, aside from one. This library, the Non-Enriched Sort 2 (6nM Gtf) library, had lower quality sequencing reads which may account for not finding this sequence.

The Non-Enriched Sort 3 Gate B library contained 22 unique sequences (**Figure 13b**), with one member occurring 23%, and an additional three each representing 11% of the sampling. The Enriched Library Sort 3 had twenty unique sequences (**Figure 13c**), with three sequences each representing more than 18% of the library. This library contained

many stop codons, a result of insertions and deletions upstream of the variable region. Of the sequences identified, only seven had mutations only in the intended positions. An activity assessment of biotin-labeled Gtf material was performed on two separate biotinylated batches, and was found to be inactivated. The original material used on the magnetic beads was depleted, and not tested for activity; however, no attempts to biotinylate the Gtf yielded active protein.

Though there was no apparent bias seen within the quality assessment of the naïve library, a potential binding motif did emerge for select mutated positions in the Non-Enriched Sort 3 Gate B population (**Figure 14**). Residues on the second alpha helix, notably mutation positions 8-12, appear to favor aromatic amino acids surrounding a charged residue.

Binding and Inhibition Assessment of Candidate Affibody Variants

Sixteen of the affibody variants identified in the sort 3 libraries (**Table 3**) were cloned into *E. coli* and analyzed for expression. Of these sixteen, nine were expressed and purified successfully. The dominant sequence with the highest library representation, called Variant A, was analyzed for Gtf binding using several methods. A dissociation constant (Kd) of between 10-250 nM was estimated using an Octet HTX System biosensor (**Figure 16**). Curiously, the raw binding data obtained were manifested in an inverted association and disassociation curves. The negative wavelength shift is indicative of extremely large particles (29), such as phage and liposomes, and is an unexpected result considering the size of the Gtf in comparison to these large particles. It could be speculated that this protein exists as a large aggregate, or that residual large dextran molecules from

the purification are binding Gtf units into aggregates that impact signal production in this assay method, however, more studies should be undertaken to explain this surprising phenomenon. Though a rough estimation, this analysis confirms Variant A as a true, non-specific Gtf binder in the tens to hundreds of nanomolar range. This approximation agrees with the apparent affinities observed for the yeast-displayed variants in this population, as well as consistent with the selection conditions for which Variant A was isolated. This affibody variant could serve as an adequate backbone for further engineering to obtain an even tighter Gtf binder.

Each of the nine successfully expressed affibodies was tested for Gtf inhibition at several conditions. No obvious inhibition of the Gtf was observed for these expressed variants in end-point assays. Time-resolved inhibition assays would be more sensitive, particularly for the weak inhibitors expected from the first round of Gtf binding selections, though project time constrains did not allow for the development of this kind of kinetic analysis. Many factors, including the pH of the reaction and pre-incubation time length, may have affected the binding capacity or stability of the variant. Furthermore, a larger assessment of the identified sequences for both binding capacity and inhibition could be more foretelling of the potential to identify an inhibitor within an affibody library. It is of note that many of the affibody variants were not expressed successfully. However, stability issues may arise for these new variants expressed in solution, as compared to being stably tethered to the surface of a yeast cell. Additionally, many of the identified sequences had additional backbone mutations which may have proved fortuitous for folding, expression, or binding. No obvious trend was observed when considering off target mutations, and for

the scope of this project, these were discounted to avoid additional complexity in validation of the identified variants.

V. Future Directions

Of the numerous sequences identified as potential Gtf binders from the generated affibody libraries, only ten were assessed for Gtf inhibition and only one for Gtf binding. The remaining sequences identified from each library still maintain a high potential for binding and inhibiting Gtf. Improvement of expression levels, particularly of those under detectable levels, could be undertaken. Additionally, the binding assays used also require additional optimization for more precise binding measurements. Difficulties surrounding the potential aggregation of the Gtf in the binding assessment could be avoided by immobilizing the Gtf and analyzing the binding to the affibody in solution. Improved binding analysis could provide the validation of a Gtf-binding motif which would be beneficial in subsequent engineering campaigns.

The Gtf selected for screening (a dextransucrase from *W. confusa)* is only one of many relevant Gtf targets for inhibiton. The naïve affibody variant library could be probed for additional binders and potential inhibitors of Gtfs from *Leuconostoc* and *Streptococcus*. Though a Gtf inhibitor has not yet been identified from among the isolated variants, the discovery of an inhibitor would warrant biophysical characterization and potential application screening. Improvements to the affibody inhibitor in the form of expression, activity, or stability could be accomplished by designing an error prone PCR library of the variant, as well as site saturation libraries of identified residues of interest. This type of affinity maturation, also employed in monoclonal antibody development, is relatively straightforward.

Strategically, the information gained from this research could aid in the discovery of an inhibitor by improving the experimental design of this approach. For instance, screening for binders with a truncated version of an active Gtf could minimize the chances of isolating non-inhibitors. By minimizing the number of antigen recognition sites away from the active site, more relevant binding sequences could be uncovered. This is a stringent approach, but may increase the chances of success. Alternatively, other small protein binders, such as monobodies, anticalins, and designed ankyrin repeat proteins (DARPs) could be utilized in a similar engineering campaign in the search for a Gtf inhibitor (30).

The exciting capacity of the affibody library for novel binding activities and potential for inhibitor discovery is not limited to Gtfs. Other biologically relevant targets should also be examined. Considering the ease of assembly and rapid screening possibilities, this approach is a powerful alternative to antibody discovery and design. As more information is acquired for this small protein, it is conceivable a model could be constructed to rationally predict variants for desired binding properties. This modeling tool could be invaluable within the medical and biotechnological fields.

VI. Conclusion

An affibody yeast display library was successfully assembled with high fitness and probed for Gtf binders. Analysis of a single variant has validated the approach to locate Gtf binders within the range of ten to hundreds of nanomolar binders. The examination of a small sampling of variants within additional binding populations did not unearth an inhibitor; however, it is still plausible that this approach will lead to such a molecule. This effort yielded an abundence of valuable insight for future strategies, which can be employed in a variety of unique directions, each as exciting and full of possibility as the next.

VII. References

1. Breton C, Fournel-Gigleux S, Palcic MM. Recent structures, evolution and mechanisms of glycosyltransferases. Curr Opin Struct Biol. 2012 Oct;22(5):540-9.

2. McArthur JB, Chen X. Glycosyltransferase engineering for carbohydrate synthesis. Biochem Soc Trans. 2016 Feb;44(1):129-42.

3. Amari M, Arango LFG, Gabriel V, Robert H, Morel S, Moulis C, et al. Characterization of a novel dextransucrase from Weissella confusa isolated from sourdough. Appl Microbiol Biotechnol. 2013 /06/01;97(12):5413-22.

4. Kajala I, Shi Q, Nyyssölä A, Maina NH, Hou Y, Katina K, et al. Cloning and characterization of a Weissella confusa dextransucrase and its application in high fibre baking. PLoS ONE. 2015;10(1):e0116418.

5. Ren Z, Chen L, Li J, Li Y. Inhibition of Streptococcus mutans polysaccharide synthesis by molecules targeting glycosyltransferase activity. Journal of Oral Microbiology. 2016;8:31095.

6. Simpson CL, Giffard PM, Jacques NA. Streptococcus salivarius ATCC 25975 possesses at least two genes coding for primer-independent glucosyltransferases. Infect Immun. 1995 Feb;63(2):609-21.

7. Kopec LK, Vacca Smith AM, Wunder D, Ng-Evans L, Bowen WH. Influence of antibody on the structure of glucans. Caries Res. 2002 Mar-Apr;36(2):108-15.

8. Zhang Q, Nijampatnam B, Hua Z, Nguyen T, Zou J, Cai X, et al. Open. Open. 2017 /07/20;7(1):5974.

9. Roychoudhury R, Pohl NL. New structures, chemical functions, and inhibitors for glycosyltransferases. Current Opinion in Chemical Biology. 2010;14(2):168-73.

10. Wang Y, Chan F, Sun N, Lui H, So P, Yan S, et al. Structure-based Design, Synthesis, and Biological Evaluation of Isatin Derivatives as Potential Glycosyltransferase Inhibitors. Chemical Biology & Drug Design. 2014 Dec;84(6):685-96.

11. Zhi Ren, Lulu Chen, Jiyao Li, Yuqing Li. Inhibition of Streptococcus mutans polysaccharide synthesis by molecules targeting glycosyltransferase activity. Journal of Oral Microbiology. 2016 Apr 1,;8:1-9.

12. Sosič I, Anderluh M, Sova M, Gobec M, Mlinarič Raščan I, Derouaux A, et al. Structure-Activity Relationships of Novel Tryptamine-Based Inhibitors of Bacterial Transglycosylase. Journal of medicinal chemistry. 2015 Dec 24,;58(24):9712.

13. Wang S, Cuesta-Seijo JA, Striebeck A, Lafont D, Palcic MM, Vidal S. Design of Glycosyltransferase Inhibitors: Serine Analogues as Pyrophosphate Surrogates? ChemPlusChem. 2015 Oct;80(10):1525-32.

14. Pesnot T, Palcic MM, Jørgensen R, Wagner GK. Structural and mechanistic basis for a new mode of glycosyltransferase inhibition. Nature Chemical Biology. 2010 May;6(5):321-3.

15. Wright WG, Thelwell C, Svensson B, Russell RRB. Inhibition of catalytic and glucan-binding activities of a streptococcal GTF forming insoluble glucans. Caries Research. 2002 Sep;36(5):353-9.

16. Case BA, Hackel BJ. Synthetic and natural consensus design for engineering charge within an affibody targeting epidermal growth factor receptor. Biotechnol Bioeng. 2016 August 1,;113(8):1628-38.

17. Mouratou B, Béhar G, Pecorari F. Artificial Affinity Proteins as Ligands of Immunoglobulins. Biomolecules. 2015 -01-30;5(1):60-75.

18. Justino CIL, Duarte AC, Rocha-Santos TAP. Analytical applications of affibodies. TrAC Trends in Analytical Chemistry. 2015 February 1,;65:73-82.

19. Nygren P. Alternative binding proteins: affibody binding proteins developed from a small three-helix bundle scaffold. FEBS J. 2008 Jun;275(11):2668-76.

20. Pacific Northwest National Laboratory. 22. Yeast Display scFv Antibody Library User's Manual. 2014.

21. Huang R, Kiss M, Batonick M, Weiner M, Kay B. Generating Recombinant Antibodies to Membrane Proteins through Phage Display. Antibodies. 2016 May 2,;5(2):11.

22. Georgescu R, Bandara G, Sun L. Saturation Mutagenesis. In: Directed Evolution Library Creation. Humana Press; 2003. p. 75-83.

23. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2007;2(1):31-4.

24. Boder ET, Wittrup KD. Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol. 1997 Jun;15(6):553-7.

25. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012 Jun 15,;28(12):1647-9.

26. Crooks GE, Hon G, Chandonia J, Brenner SE. WebLogo: a sequence logo generator. Genome Research. 2004 Jun;14(6):1188-90.

27. Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. Nucleic Acids Research. 1990 Oct 25,;18(20):6097-100.

28. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. John M. Walker (ed): The Proteomics Protocols Handbook; Humana Press; 2005.

29. Forte Bio Inc. Octet System Data Analysis User Guide. Release 7.1. 2011.

30. Gilbreth RN, Koide S. Structural insights for engineering binding proteins based on non-antibody scaffolds. Current Opinion in Structural Biology. 2012 August 1;22(4):413-20.

31. Placzek S, Schomburg I, Chang A, Jeske L, Ulbrich M, Tillack J, et al. BRENDA in 2017: new perspectives and new tools in BRENDA. Nucleic Acids Res. 2017 /01/04;45(D1):D388.

32. Tashiro M, Tejero R, Zimmerman DE, Celda B, Nilsson B, Montelione GT. High-resolution solution NMR structure of the Z domain of staphylococcal protein A. J Mol Biol. 1997 Oct 03;272(4):573-90.

33. Schrödinger L. The PyMOL Molecular Graphics System. ;Version 1.8.

VIII. Tables

Library	Good Clones	Unique Sequences	seq >10 % library
NE sort 2, 3nM	80	14	2
NE sort 2, 6nM	69	26	3
E sort 2	92	34	1
NE sort 3 Gate A	93	3	1
NE sort 3 Gate B	94	22	4
E sort 3	96	20	3

Table 1. Summary FACS Sorted Libraries Sequence Analysis

25 Shared Sequences		
# Clones	Sequence	Library
4	EYSHAMAEIPQLPNLNSWQRVAFIWSLV	NE sort 2, 3nM
1		NE sort 2, 6nM
5	EYQVAMVEIPHLPNLNIYQQQAFIWSLW	NE sort 2, 3nM
10		NE sort 2, 6nM
20		NE sort 3 Gate B
1	EYQNAAAEIFILPNLNYYQKYAFIWSLD	NE sort 2, 6nM
1		NE sort 3 Gate B
6	EYPYASEIYLLPNLNLYQKYAFIFSLP	NE sort 2, 6nM
11		NE sort 3 Gate B
2	EWYQALVEIGILPNLNLYQNWAFIYSLT	NE sort 2, 6nM
5		NE sort 3 Gate B
15	ERNQAWIEIPWLPNLNPWQRWAFIYSLQ	NE sort 2, 6nM
11		NE sort 3 Gate B
1	EREFASMEIRWLPNLNIYQKYAFIYSLM	NE sort 2, 6nM
2		NE sort 3 Gate B
2	EQWLASFEISQLPNLNRNQVQAFINSLQ	NE sort 2, 3nM
4		E sort 2
5		NE sort 3 Gate B
91		NE sort 3 Gate A
1		E sort 3
1	EPINAYYEIITLPNLNPYQRYAFIYSLQ	NE sort 2, 3nM
1		NE sort 3 Gate B
34	EPFAAQAEIDWLPNLNRWQQLAFIWSLH	NE sort 2, 3nM
11		NE sort 3 Gate B
1		NE sort 3 Gate A
1	ENGYAMFEINMLPNLNIYQKWAFIFSLQ	NE sort 2, 3nM
1		NE sort 2, 6nM
3		NE sort 3 Gate B
15	EMMYARPEIMFLPNLNPWQKVAFIYSLQ	NE sort 2, 3nM
1		NE sort 2, 6nM
2		NE sort 3 Gate B
5	EIFNARSEIWRLPNLNIFQAYAFILSLE	NE sort 2, 3nM
1		NE sort 2, 6nM
5	EEYMASHEIWRLPNLNIYQQYAFITSLQ	NE sort 2, 3nM
1		NE sort 2, 6nM
7		NE sort 3 Gate B
2	EERDAKLEIYILPNLNPYQKYAFIWSLE	NE sort 2, 6nM
5		NE sort 3 Gate B
1	EEPIALPEILLLPNLNIWQRWAFIWSLN	NE sort 2, 6nM
1		NE sort 3 Gate B
2	EAQMADSEIRKLPNLNIYQRYAFIWSLV	NE sort 2, 3nM
1		NE sort 2, 6nM
1		NE sort 3 Gate B
15	ETWLAGFEITQLPNLNNVQVRAFIISLK	E sort 2
4		E sort 3
1	EEMLATTEIQLLPNLNVFQVTAFIGSLD	E sort 2
1		E sort 3
5	EMVLARMGPI*TFIRTWPSSSH*	E sort 2
2		E sort 3
1	EMRHAAMEILHFAQFERIPA*RLHPVTA	E sort 2
1		E sort 3
3	EMQNAQMEILGLPNLNIRQFEAFIKSLL	E sort 2
1		E sort 3
1	EMPMAPKEILRFAQFELVPAAGLHRFTA	E sort 2
21		E sort 3
6	EIQPARLEIVWFAQFEPAAVGCLHHFTV	E sort 2
22		E sort 3
3	EINLAQMEIQVFAQFEPLPV*RLHHVTA	E sort 2
18		E sort 3

Table 2. Sequence Analysis of Six Gtf Binding Libraries

A summary of the 25 shared sequences, located in two or more libraries analyzed, containing the libraries of origin and a heat map of library representation.

ID	Sequence	Library	Expression	Inhibition
Wt	VDNKFNKEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKLNDAQAPK		Y	N
A	VDNKFNKEQWLASFEISQLPNLNRNQVQAFINSLQDDPSQSANLLAEAKKLNDAQAPK	1	Y	N
B	VDNKFNKEPFAAQAEIDWLPNLNRWQQLAFIWSLHDDPSQSANLLAEAKKLNDAQAPK	1	Y	N
C	VDNKFNKEYQVAMVEIPHLPNLNIYQQQAFIWSLWDDPSQSANLLAEAKKLNDAQAPK	2	Y	N
D	VDNKFNKEPFAAQAEIDWLPNLNRWQQLAFIWSLHDDPSQSANLLAEAKKLNDAQAPK	2	N	–
E	VDNKFNKERNQAWIEIPWLPNLNPWQRWAFIYSLQDDPSQSANLLAEAKKLNDAQAPK	2	N	–
F	VDNKFNKEEYMASHEIWRLPNLNIYQQYAFITSLQDDPSQSANLLAEAKKLNDAQAPK	2	Y	N
G	VDNKFNKEERDAKLEIYILPNLNPYQKYAFIWSLEDDPSQSANLLAEAKKLNDAQAPK	2	Y	N
H	VDNKFNKEWYQALVEIGILPNLNLYQNWAFIYSLTDDPSQSANLLAEAKKLNDAQAPK	2	N	–
I	VDNKFNKEMMYARPEIMFLPNLNPWQKVAFIYSLQDDPSQSANLLAEAKKLNDAQAPK	2	N	–
J	VDNKFNKEREFASMEIRWLPNLNIYQKYAFIYSLMDDPSQSANLLAEAKKLNDAQAPK	2	N	–
K	**VDNKFNKEAQMADSEIRKLPNLNIYQRYAFIWSLVDDPSQSANLLAEAKKLNDAQAPK**	2	Y	N
L	VDNKFNKEVKWATLEILPLPNLNWLQDKAFIASLEDDPSQSANLLAEAKKLNDAQAPK	3	Y	N
M	VDNKFNKETWLAGFEITQLPNLNNVQVRAFIISLKDDPSQSANLLAEAKKLNDAQAPK	3	N	–
N	VDNKFNKEYYMAGPEILRLPNLNRTQRMAFIQSLFDDPSQSANLLAEAKKLNDAQAPK	3	Y	N
O	VDNKFNKEMQNAQMEILGLPNLNIRQFEAFIKSLLDDPSQSANLLAEAKKLNDAQAPK	3	Y	N
P	VDNKFNKEEMLATTEIQLLPNLNVFQVTAFIGSLDDPSQSANLLAEAKKLNDAQAPK	3	N	–

Table 3. Sequences Selected for Expression, Binding and Inhibition Analysis

The sixteen affibody variants which were selected for expression and inhibition studies. Seven of the sixteen were not expressed successfully, and the remainder did not demonstrate Gtf inhibition in an end-point assay. The library associated with each sequence is listed, with the Non-Enriched Gate A (Sort 3) library annotated as library 1, the Non-Enriched Gate B (Sort 3) as library 2, and Enriched (Sort 3) as library 3. Mutated residues are highlighted in purple.

IX. Figures

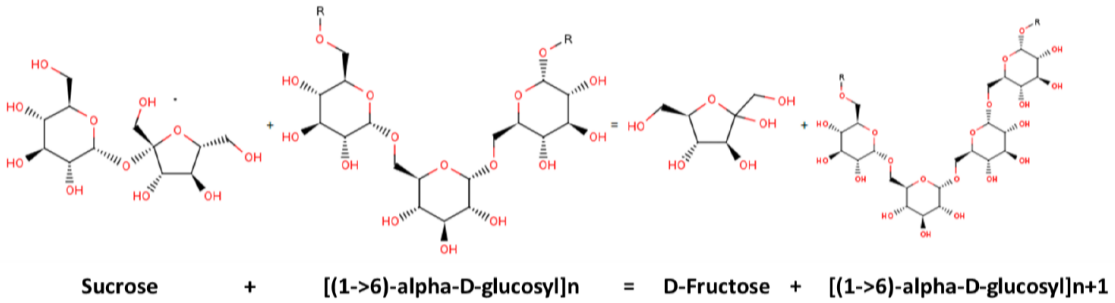

Figure 1. Dextransucrase Reaction (Adapted from 31)

Dextransucrase are glucosyltransferases (Gtfs) which catalyze the transfer of a sugar residue to an accepting polysaccharide chain in an alpha 1,6-linkage.

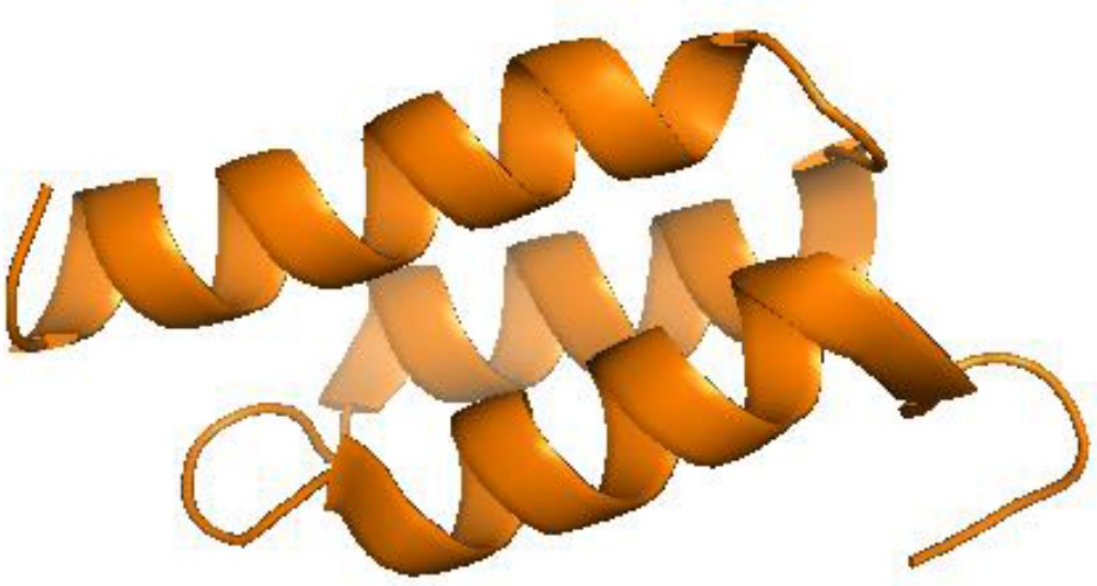

Figure 2. Z-domain of Staphylococcal Protein A (pdb 2spz) (32)

The affibody wild-type (Wt) backbone commonly used for protein engineering new binders. Comprised of just three connected alpha helices, this small protein exhibits desirable qualities, such as high stability without disulfide bonds. The high tolerance to

mutations on thirteen identified residue positions on two of the three alpha helices offers an increased potential for developing new binding properties. Image generated using PyMOL(33).

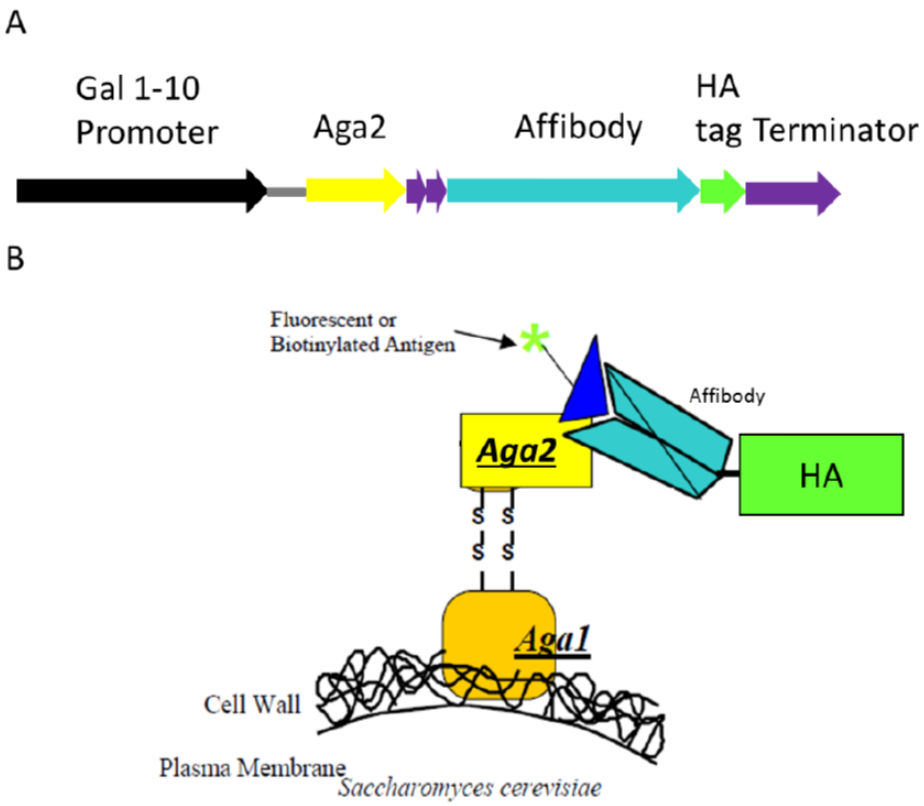

Figure 3. Construction of a Yeast Display Affibody Library

Illustration representing the A) Surface expression construct B) The affibody Aga2 fusion protein surface expression system (Illustration adapted from 20). The assembled construct contains a galactose inducible system used to initiate expression and localization of affibody variant to the extracellular surface of the yeast by covalent binding of the reproductive proteins Aga1 and Aga2. The affibody additionally contains an HA epitope for expression assessment. Anchored to the surface of yeast, the affibody can bind to extracellular targets.

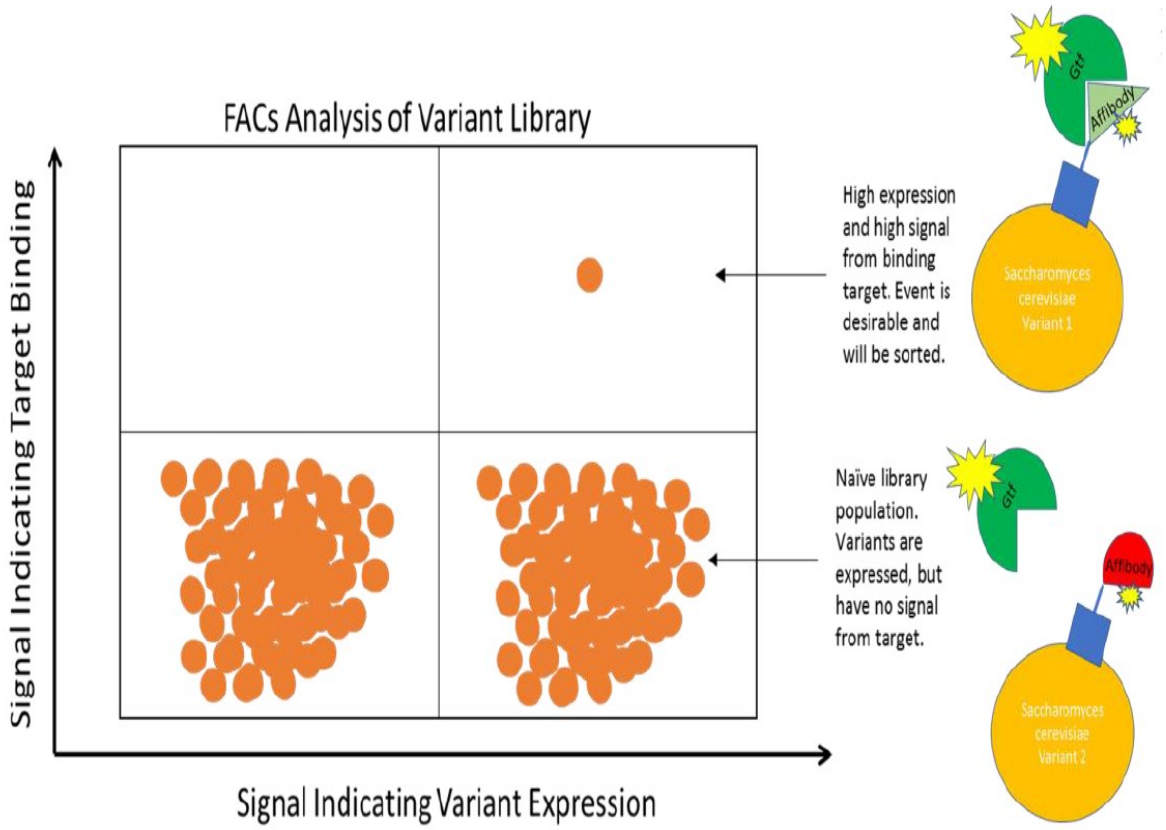

Figure 4. FACS Sorting Yeast Display Libraries

A simplified illustration of FACs analysis used to sort a library of variants. Each dot represents a single cell event. Variants are displayed on the surface of the yeast with a fluorescent tag to detect expression levels. To be sorted out of the sample solution, the yeast cell must have a high expression level in addition to a signal from the bound fluorescent target. (Image adapted from 20)

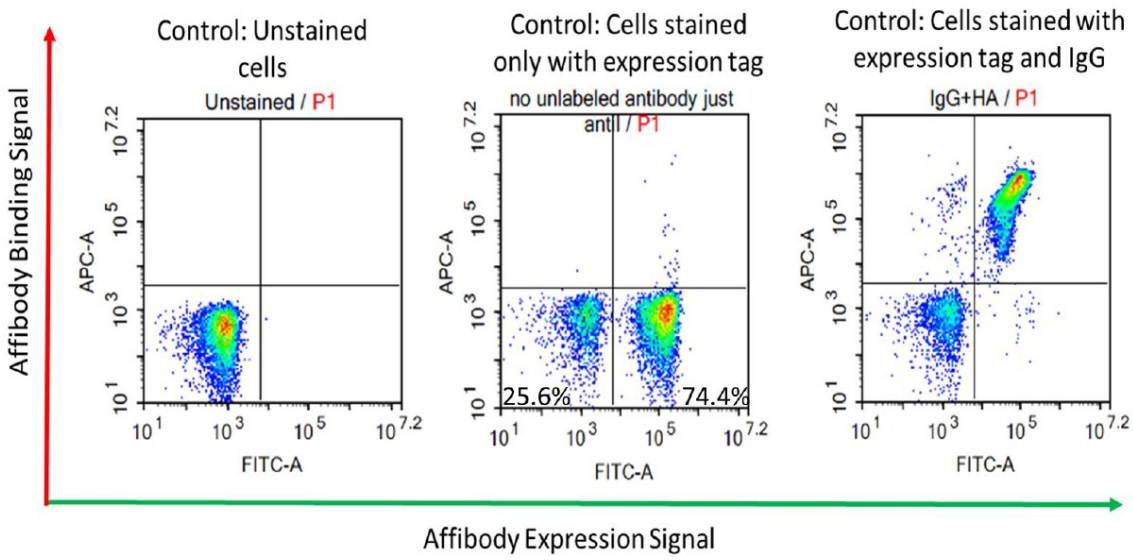

Figure 5. **Analyzing Wt Yeast Display Library**

Flow cytometry density plots illustrating Wt affibody expression (x-axis) and IgG binding (y-axis) within the yeast display system. No background in either fluorescence channel was detected for unstained cells. The Wt library exhibited high expression levels in the yeast display system. Nearly all of the expressed population were shown to have bound to IgG, indicating proper folding of the displayed protein.

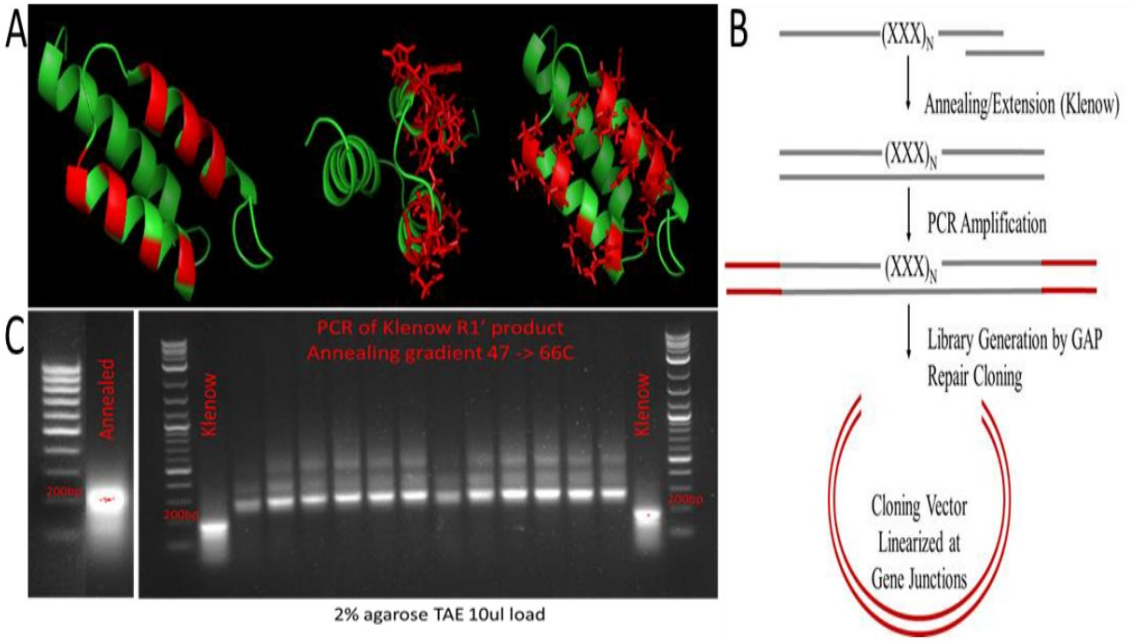

Figure 6. Design and Cloning Strategy of an Affibody Variant Library

A) The structure of the Wt affibody (pdb 2spz) at various angles, with the thirteen amino acids targeted for saturation mutagenesis highlighted in red. Image generated using PyMOL (33). B) The cloning strategy employed to assemble the affibody variant library. Single stranded oligonucleotides containing variable regions were annealed and extended in a Klenow reaction, and the resulting material was further amplified by PCR. The final material was transformed with the linearized cloning vector into *S. cerevisiae* to be generated in the host by gap repair. C) Ethidium bromide stained agarose gels illustrating the annealed, Klenow, and PCR products generated during the library cloning.

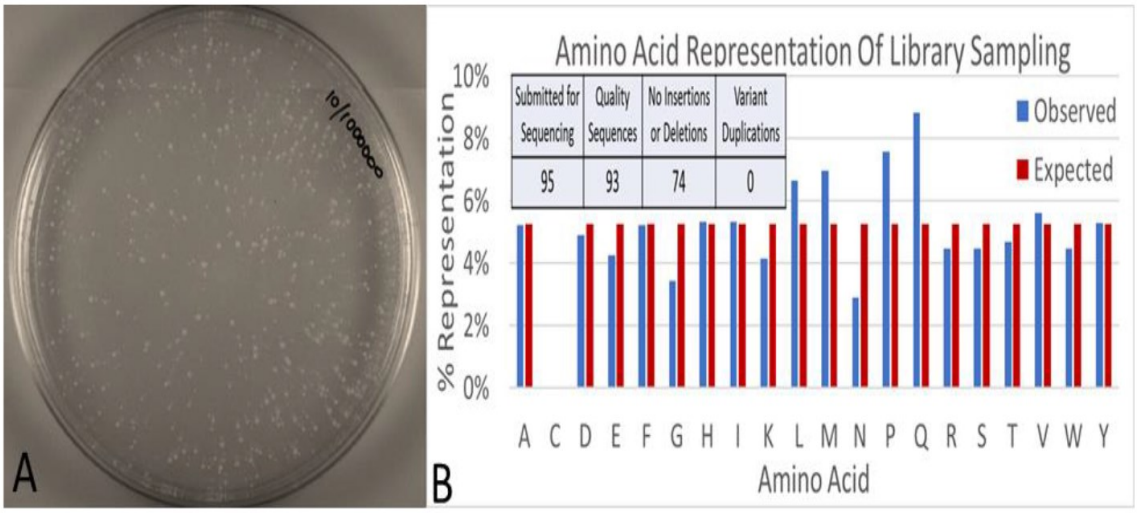

Figure 7. Diversity and Fitness Assessment of Affibody Variant Library

A) Dilution plate of transfected yeast, used to estimate diversity. B) Sequence quality and amino acid bias of a sampling of the assembled variant library.

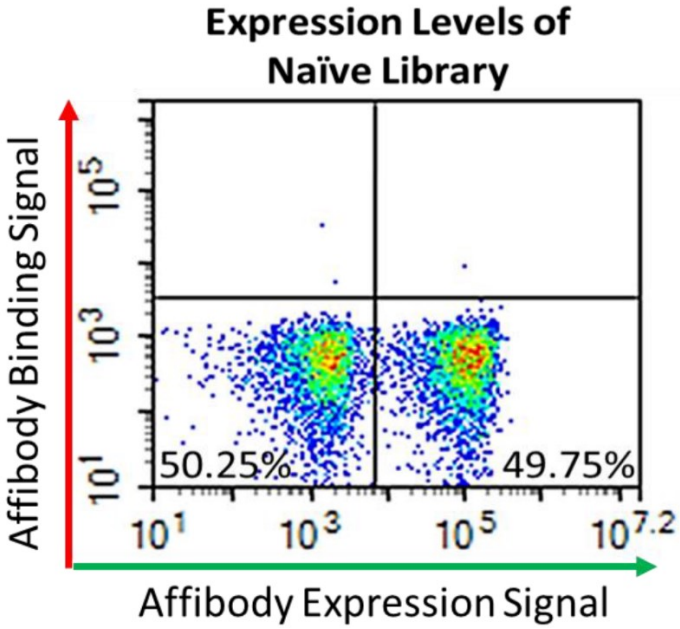

Figure 8. **Expression and IgG Binding Evaluation of Assembled Library**

Flow cytometry density plot of the naïve library expression levels (x-axis) and IgG binding (y-axis).

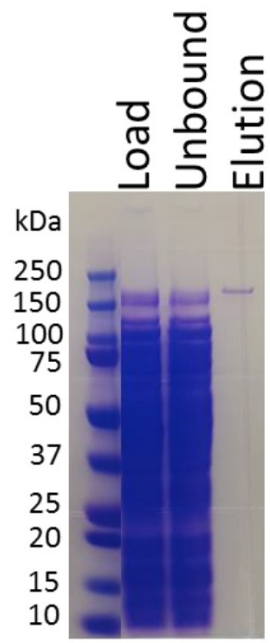

Figure 9. SDS-PAGE Analysis of Purification of Gtf from *W. confusa*

Ultrafiltrate of recombinantly expressed Gtf target was purified using Sephadex75 resin and eluting with dextran. By SDS-PAGE analysis, the protein elution appears to be of high purity.

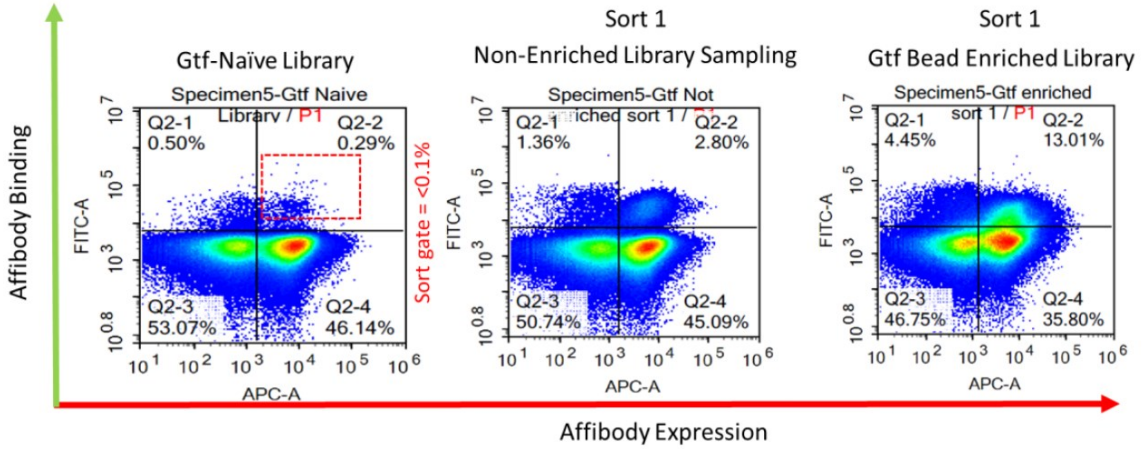

Figure 10. Analysis of FACS Sorted Libraries

Flow cytometry density plot of the library expression levels (x-axis) and Gtf binding (y-axis). Here, the Naïve (unsorted) library Gtf binding is compared to the two FACS sorted

libraries, Non-Enriched (Sort 1) and Enriched (Sort 1). An emerging population of binders can be clearly seen for both libraries.

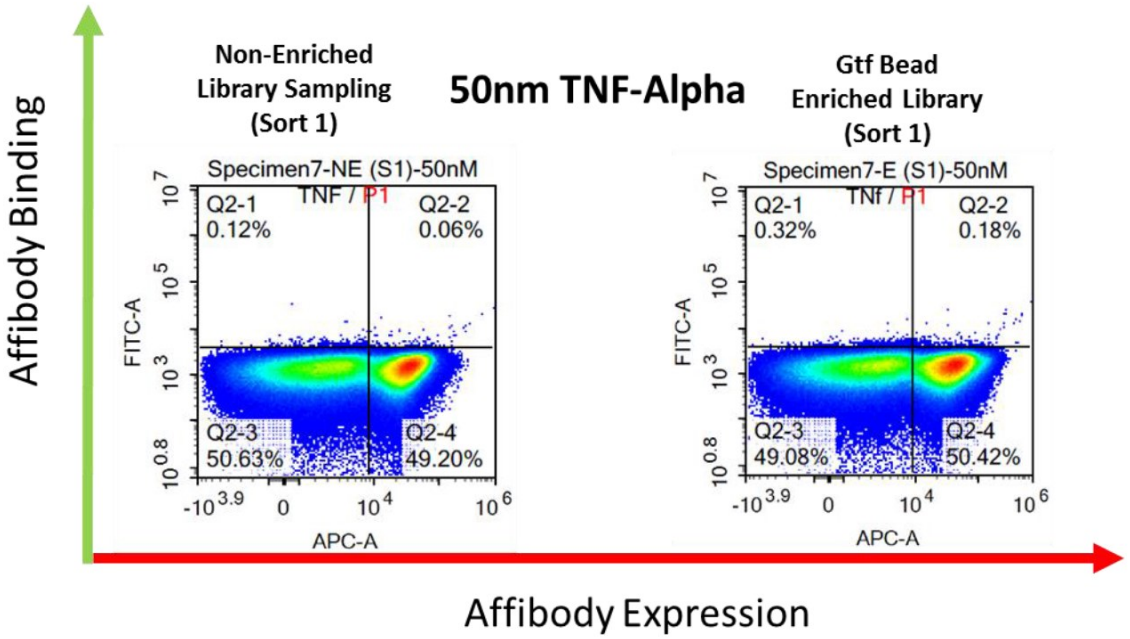

Figure 11. Off Target Library Binding Assessment

The FACS sort 1 of both the Enriched and Non-Enriched library were tested for binding against an off-target molecule, TNF-alpha (y-axis). No binding to this molecule was observed.

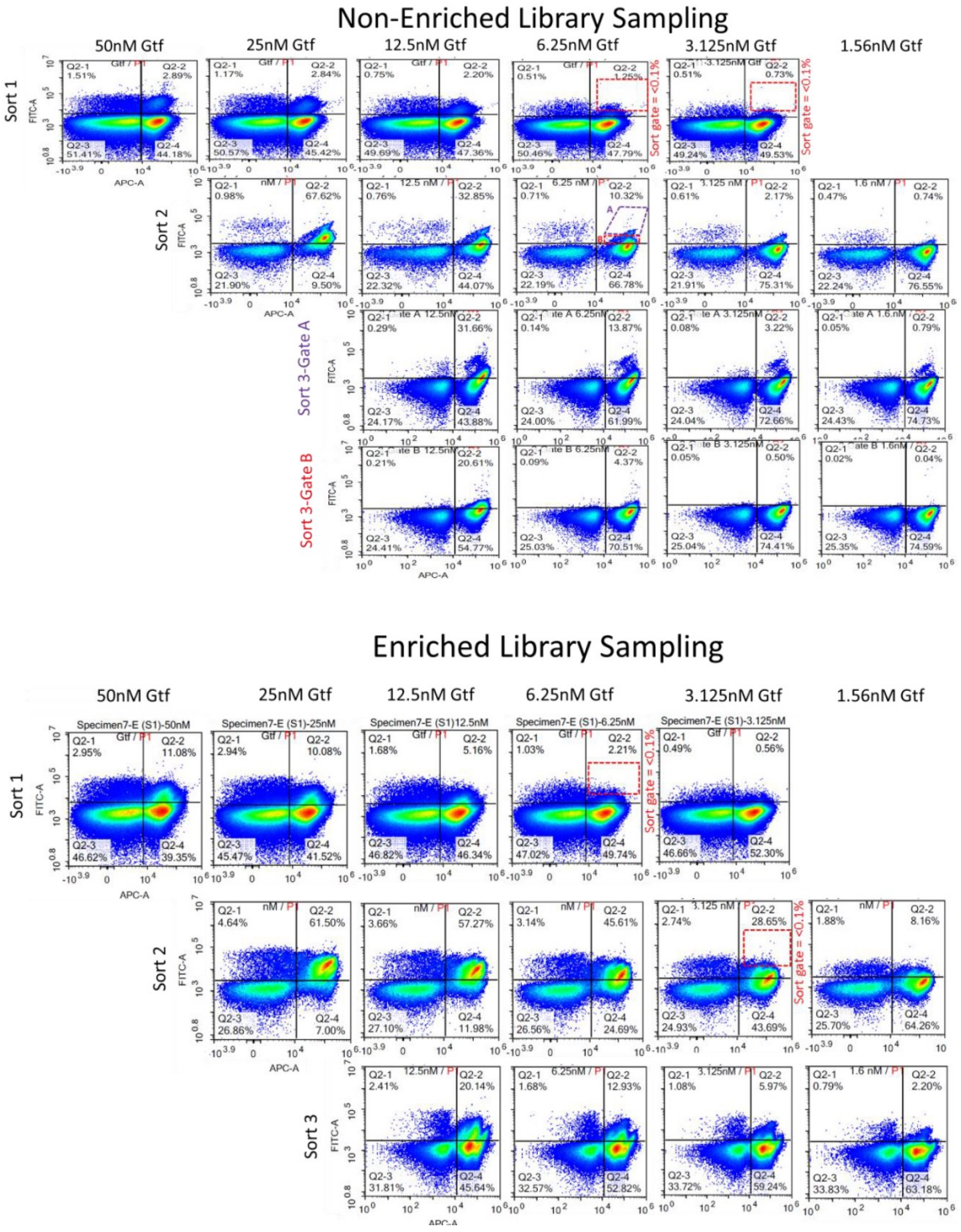

Figure 12. Gtf Binding Analysis of Library Sorts 1-3

Flow cytometry density plot of the library expression levels (x-axis) and Gtf binding (y-axis). A titration of the Gtf target can be seen from 50nM to 1.56nM. For both the Non-

Enriched and Enriched libraries, a dose dependent binding signal is observed. After each round of FACS selection, an increased population of Gtf binders is observed.

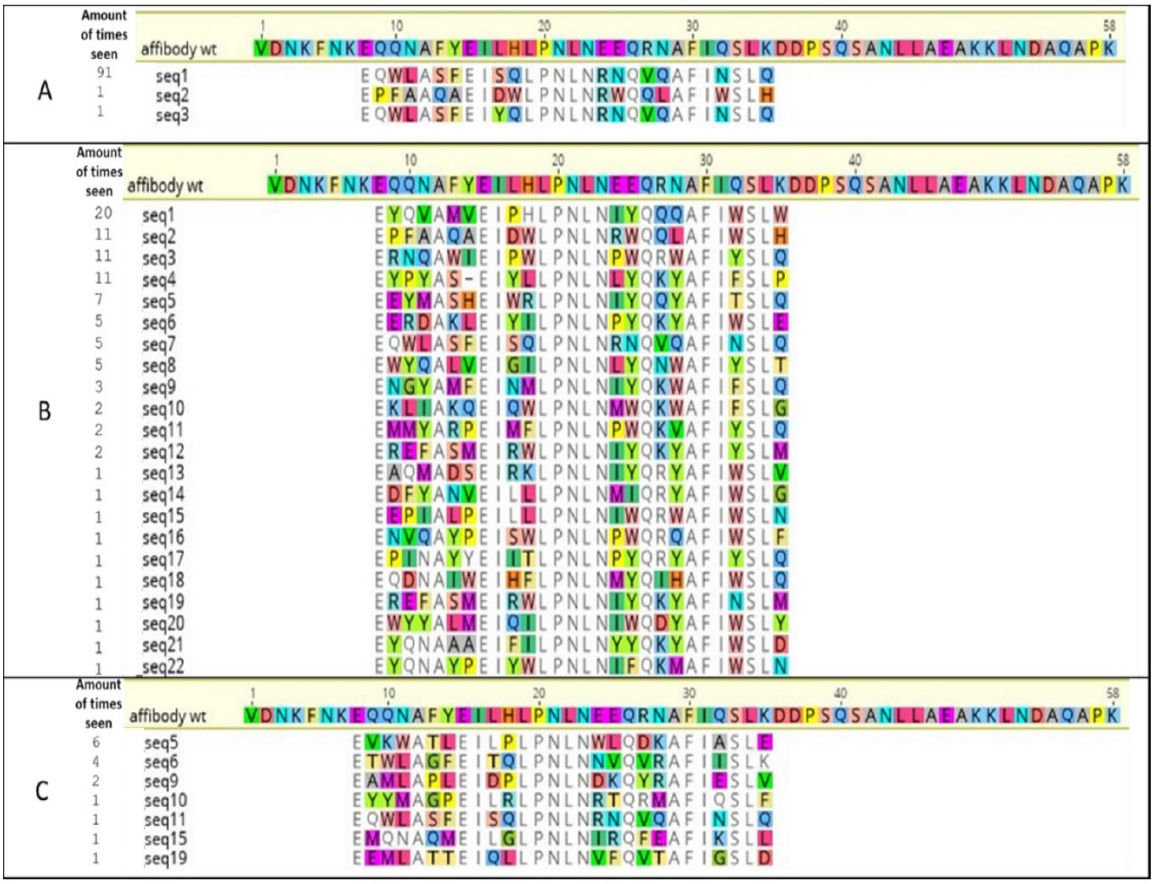

Figure 13. Library Sequence Analysis

A sampling of A) The Non-Enriched Gate A library B) The Non-Enriched Gate B library and C) The Enriched library from the last round of selection (Sort 3) were analyzed for unique sequences and their overall representation in the population.

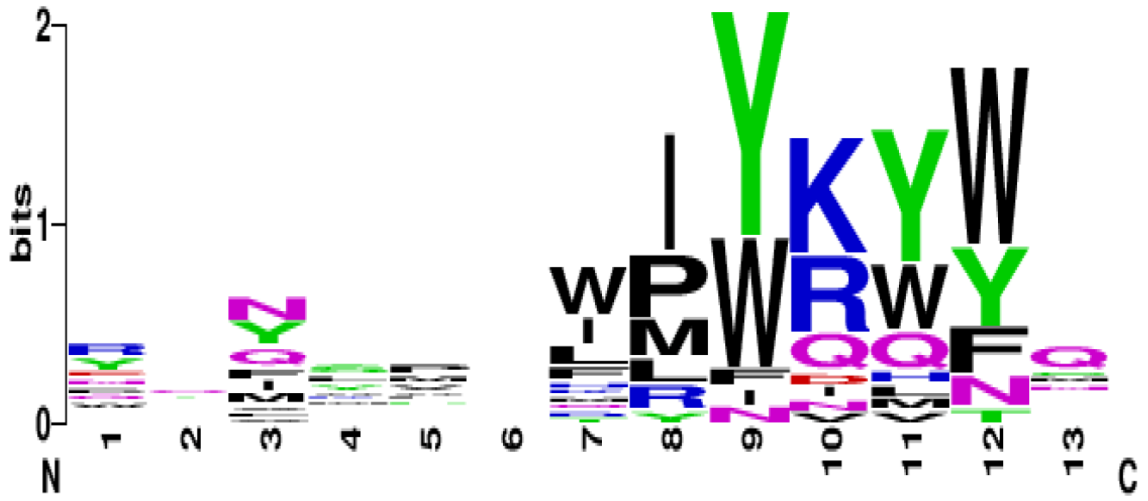

Figure 14 Potential Gtf Binding Motif

An analysis of the population of binders from the Non-Enriched Library Gate B (Sort 3) illustrates a potential emergence of a binding motif found on the second alpha helix, particularly on residues 8-12.

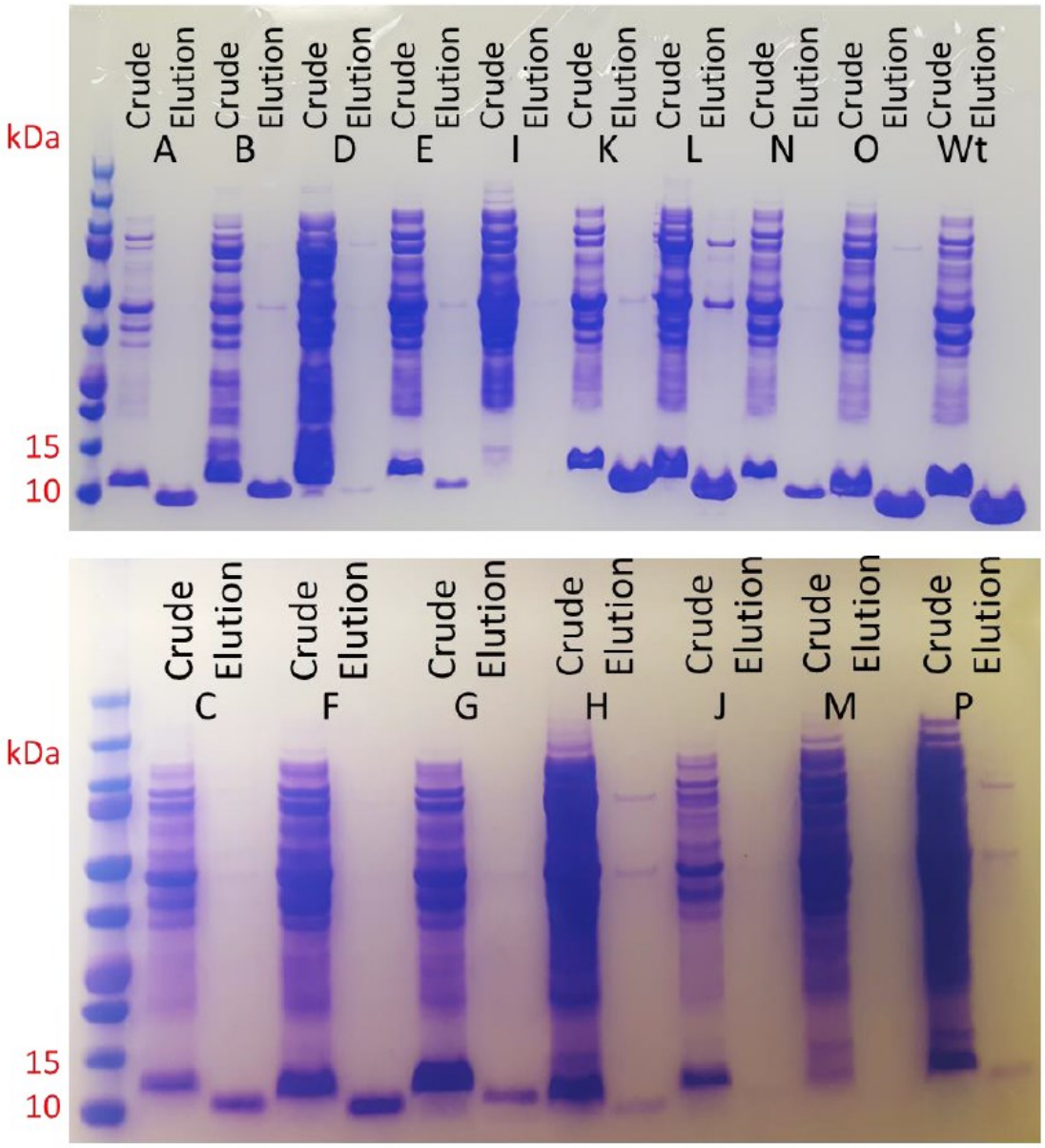

Figure 15 Expression Purification of Several Affibody Variants

Sixteen sequences identified from the Sort 3 libraries were cloned, expressed, and purified using affinity chromatography.

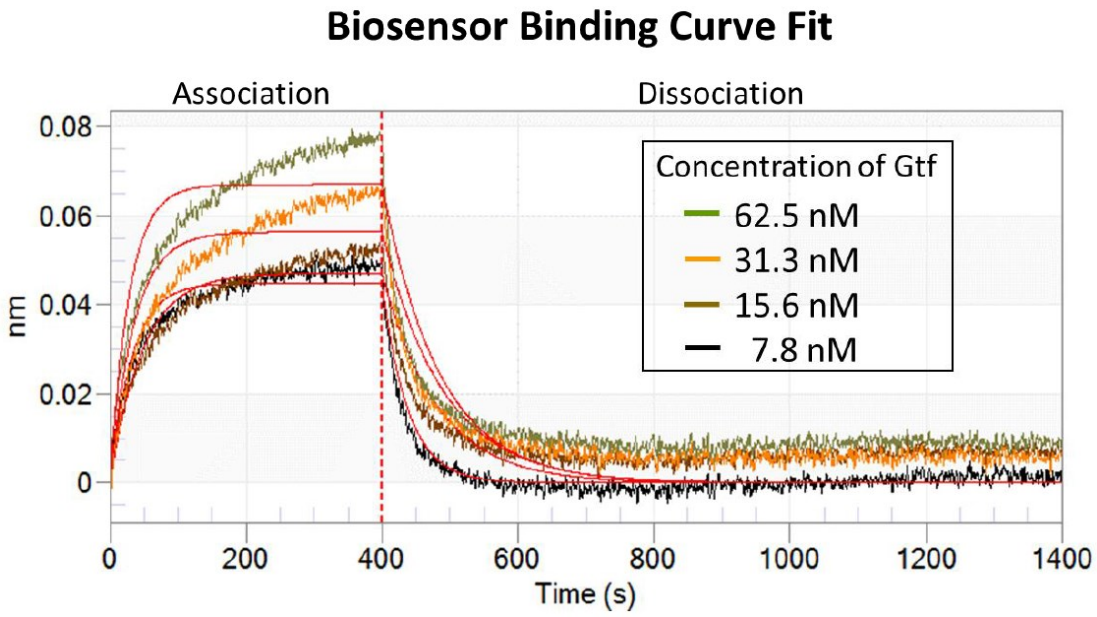

Figure 16 Gtf Binding Curve Fit of Affibody Variant A

The binding of Variant A was analyzed against several Gtf concentrations and the binding response was fit to a curve. The estimated dissociation constant was predicted to be between 10-250nM. Data acquired was inverted for proper assessment.

X. Biography

Angel D'Oliviera was born and raised in southern New Jersey. Her enthusiasm for science blossomed during her enrollment in the High School Academy of Biological and Medical Sciences. She graduated from Rowan University in 2011 with a Bachelor of Science in Biochemistry, and has since gained over 6 years of experience in the biotechnology industry working for DuPont as a Biochemist. Her diverse background includes genetic and protein discovery within the agricultural industry, as well as protein engineering for home and personal care applications.